MegaQI Covariate Analysis and Recommendations: Identification and Evaluation of Existing Quality Indicators that are Appropriate for Use in Long-Term Care Settings.

Contract No. 500-95-0062 TO #4

December 20, 2002

Prepared for
Yael Harris
Centers for Medicare and Medicaid Services
Room S3-05-26
7500 Security Boulevard
Baltimore, MD 21244-1850

Prepared by
David Kidder
Melissa Rennison
Henry Goldberg
David Warner
Barbara Bell
Louise Hadden
John Morris
Richard Jones
Vince Mor

Contents

Executive Summary

Over the past two years, the MegaQI Team Steering Committee (SC)[1] selected and validated 45 chronic care (CC) and post-acute care (PAC) quality measures (QMs) (Exhibit 1).[2] Based partly on recommendations from the SC and from the National Quality Forum (NQF), the Centers for Medicare and Medicaid Services (CMS) chose five CC and three PAC QMs for public reporting by posting on CMS' Nursing Home Compare website in November 2002. Two of these QMs (one CC and one PAC) were posted both with and without facility-level adjustment.

Some of the 45 QMs are adjusted for resident-level risk factors, or resident "covariates." Also, adjustment for some QMs includes a facility level measure (the facility admissions profile, or FAP). After selecting QMs for public reporting, CMS instructed the SC to revisit risk adjustment, focusing specifically on resident-level factors. This report presents the results of that process.

The argument for risk adjustment is simple. A well-designed system should make quality measurement fairer by adjusting for risks that facilities cannot control. However, designing a workable system is not simple. In part, this is because it is difficult to measure true, uncontrollable risk. Nonetheless, the SC believes that risk adjustment is necessary and feasible for some QMs. We chose to base risk adjustment on statistical modeling, to allow many covariates to define multiple levels of risk.

We approached this task systematically, in a series of steps that combined analysis, review and decision-making on the proposed covariates by the MegaQI Team's Steering Committee.

Step 1: Selecting the Initial List of Covariates. In September 2002, the SC met to select covariates for initial testing. We began with a long list of candidate covariates, applying several criteria to narrow the list. In addition to standards of statistical correlation, we selected covariates 1) that SC clinicians and researchers believed to be good candidate measures of risk, 2) that the SC believed could not be easily "gamed" by nursing facilities to improve their scores, 3) that facilities were unlikely to have "caused" by their previous actions and 4) that could be constructed from the Minimum Data Set (MDS). As shown in Exhibit 2, we selected:

- the 30 "original" covariates already included in specifications for the 45 CC and PAC QMs;
- indices and scales based on the Resource Utilization Group-III (RUG-III) Case Mix system;
- the Nursing Severity Index (NSI);
- two new indices, the Personal Severity Index (PSI) and a Cardio/Pulmonary Impairment Severity Scale; and
- 12 diagnosis indicators.

QMs were computed from data gathered in "target" resident assessments. All covariates were computed from assessments conducted prior to the target assessments

[1] MegaQI Steering Committee members are from CMS, Abt Associates, Hebrew Rehabilitation Center for Aged and Brown University.

[2] Exhibits cited here and in the main report are attached to the main report. In addition, we have provided a separate Technical Appendix with specifications for the covariates and models, and detailed statistical tests.

Step 2: Reviewing Correlations of Covariates and QMs at the Resident Level. Next, the SC explored the statistical relationship between QMs and covariates. It was agreed that if a measure of correlation between a covariate and a QM was 0.10 or larger, we would retain the covariate for further analysis.[3] Using this standard, and with continued review and discussion of clinical, behavioral and other factors, we dropped the following: 17 of the 30 original covariates; RUG-based scales for Extensive Care, Extensive Services, Rehabilitation and Special Care; the NSI; the Cardio/Pulmonary Impairment Severity Scale; and all diagnosis indicators except acute episode, Alzheimer's disease, other dementia, and hip fracture. The SC combined Alzheimer's and other dementia in one covariate in the final specifications.

Step 3:Testing Covariates in Resident-Level Prediction Model. In Step 3, our objective was to design resident-level models that related specific covariates to the appropriate QMs. With these models, we could show how resident risks related to the prevalence of resident problems. For example, we could predict how much more prevalent bowel incontinence is likely to be among residents with three of four possible risk factors, compared to residents with only two risk factors. To build prediction models, we used MDS data from a sample of all Quarter 2 2002 nursing home residents. We studied how closely each covariate correlated with its associated QM, alone and together with other covariates. After examining several measures of how well different combinations of covariates "fit" the data, we selected final lists of covariates for each QM model. Then, we retested our models using data from five new samples of residents. The models performed well in the retest. The SC approved covariates and models developed in Step 3 for further tests at the facility level in Step 4.

Step 4: Analyzing Facility-Level Risk Adjustment. In Step 4, we used the resident-level prediction models to compute risk adjusted QMs at the facility level and to study the effects of adjustment on QM scores and rankings of nursing facilities. In the adjustment process, each facility's observed and predicted scores are combined with the national mean, to estimate what that facility's QM would be if it faced the national "average" mix of resident risks. We tested the new adjustment models against previous models (based on the FAP, and based on the original resident-level covariates). We were particularly interested to study how effectively adjustment "targets" facilities ranked highest and lowest both in observed QMs and in measures of resident risk, measured by the RUG-III Case Mix Index.

In general, facility-level analyses of adjustment based on the new covariate models showed improvement over the FAP and/or original covariate models, particularly in targeting effectiveness. As expected, we found variation in comparative performance. Some new models achieved solid improvement measured in almost all comparisons with alternatives, most performed well in some comparisons, and a few showed little or no relative improvement.

Recommendations. At the conclusion of this process, the SC recommended resident-level covariates and models, shown in Exhibit 3, that represent a significant departure from the original QM specifications. Our analyses suggest that there may be room for further exploration of the role of nursing facility-level measures in QM adjustment. But we consider the resident-level covariates that we recommend to be valid and accurate for adjusting publicly reported nursing facility QMs

[3] In fact, the SC retained some covariates that members supported strongly on clinical or behavioral grounds, but that had correlations slightly lower than 0.10.

Background

Over the past two years, the MegaQI Team Steering Committee (SC), including investigators from Abt Associates and colleagues at Hebrew Rehabilitation Center for Aged (HRCA), Brown University, and the Centers for Medicare and Medicaid Services (CMS), developed specifications for 45 chronic care (CC) and post-acute care (PAC) nursing facility quality measures (QMs). (See Exhibit 1 for a list of all QMs).[4] The National Quality Forum (NQF) reviewed the SC's methodology and recommended a short list of measures for immediate public reporting. With input from a steering committee convened by the NQF and other experts, CMS selected eight of the 45 (five CC and three PAC QMs) for posting on CMS' Nursing Home Compare website in November 2002. Two of these QMs (one CC and one PAC) were posted both with and without facility-level adjustment.

Specifications for most QMs called for adjustment based both on covariates that measure resident-level risk and, for many QMs, facility-level measures (the facility admissions profile, or FAP). Since preliminary risk adjustment analyses had been accelerated to meet public reporting requirements, CMS asked the MegaQI Team to revisit the issues of covariate selection during the fall of 2002, focusing only on resident-level covariates. Over the past two months, we have assessed covariates for QM adjustment in a more comprehensive and systematic fashion than had been possible before. This report recommends resident-level covariates for QM risk adjustment, and it describes the analyses and decision processes that we used to develop these recommendations.

Why Risk Adjust?

The argument for risk adjusting QMs is simple. The ideal risk adjustment system should calibrate QMs to better measure real differences in quality among nursing facilities. Facility A's unadjusted or "observed" QM shows a higher-than-average prevalence of problems. Facility A also cares for residents at higher-than-average risk of the problem measured by that QM. But Facility A should not be ranked equal in performance to Facility B, which has the same score but treats lower-risk residents. Risk adjustment should correct for differences in resident characteristics over which facilities have little or no control. In this example, Facility A's adjusted score, which shows what the QM would be if Facility A admitted lower risk residents (an "average" resident risk profile) should be lower than Facility B's adjusted score. Adjustment should also move scores of facilities with lower-than-average risk profiles toward the average, producing an adjusted QM higher than the unadjusted QM.

Although the argument for risk adjustment is simple, designing a workable system is not, since uncontrollable risk is difficult both to define and to measure. For example, if a resident is incontinent at admission or at a regular quarterly assessment, her risk of acquiring a pressure sore over the next three to six months may be higher than it is for a continent but otherwise similar resident. But the facility may be able to reduce this risk by treating her incontinence. If so, and if the facility can

[4] Exhibits are attached at the end of this report. In addition, we have provided a Technical Appendix with specifications for the covariates and models, and detailed statistical tests used to analyze the covariates. Note that QM names in Exhibit 1 are the expanded, consumer-friendly descriptors developed to facilitate public reporting. In the text of this document, we generally use shortened names or code names to describe specific QMs.

mitigate other residents' incontinence as well, adjusting a pressure sore QM using a covariate that measures incontinence might actually "over adjust" the QM. Over adjustment produces a QM score lower than it should be, because incontinence is treated entirely as a measure of risk, not a condition that the facility can change through appropriate care. The risk of developing pressure sores may be buried in other resident characteristics that make incontinence difficult or impossible to treat (for example, deteriorating physical or cognitive function related to the natural progression of a disease).

How much the facility is truly at risk for pressure sores and other problems is almost always debatable. Some might use these ambiguities to argue for minimal or no risk adjustment, on grounds that any adjustment system will tend to give too much leeway to nursing facilities that should address competently the problems of all residents, regardless of risk. The MegaQI Team took the position that risk adjustment is a necessary part of a fair quality measurement system, for some, though not all, QMs.

In risk adjustment, residents are assigned to risk groups . The simplest approach assigns residents to two groups, high and low risk, and reports separate QM scores for each. For example, the QM Behavior Symptoms Affecting Others, developed by the Center for Health Systems Research and Analysis (CHSRA) is defined for a high risk group (all residents with specific, relevant functional and diagnostic evidence of behavior problems) and a low risk group (residents without these indicators). The SC adopted a different method, using statistical estimation techniques to define multiple risk groups, measured by one or more resident-level covariates. We determined a resident's membership in a risk group by the presence or absence of certain risk factors (for example, dependence in toileting) or, for some measures, the level of risk captured in a multiple-value scale (for example, six levels of increasing risk captured by the Cognitive Performance Scale). With this approach, resident risk could be measured in several dimensions (dependence in toileting and a Cognitive Performance Scale score, plus others if appropriate).

Step One: Selecting the Initial List of Covariates

In September 2002, the SC met to discuss and plan for the covariate analyses. At this meeting, the SC reviewed and selected for subsequent analysis the "original" covariates (those already identified in specifications for the 45 QMs), several new scales and indices, and some proposed new diagnosis measures. Later, we conducted statistical tests to see how closely all proposed covariates correlated with the QMs. Throughout this process, in addition to statistical evidence, we reviewed covariates using clinical and other standards.

- Any proposed covariate had to have "face validity." In the opinion of experienced SC clinicians and researchers, the covariate had to be plausible, independent of any statistical evidence, as a measure of the risk of a particular QM. The SC paid particular attention to the dangers of over adjustment. We dropped several potential covariates that met our statistical criteria for inclusion but appeared to be questionable as measures of uncontrollable risk.

- The covariate should offer minimal incentive for facilities to "game the system" (for example, by recoding MDS items or by being more selective with admissions).

- We generally rejected covariates that might have been influenced by the facility's own actions. For example, a resident with pressure sores three months ago (a potential covariate) is clearly at higher risk of having pressure sores today. But the facility may have been at least partly responsible for the resident's earlier condition.. Therefore, we considered existence of prior pressure sores to be an inappropriate covariate.

- The proposed covariate could be constructed from one or more Minimum Data Set (MDS) items.

Proposed covariates included:

- the 30 "original" covariates in the specifications for the 45 CC and PAC QMs -- these covariates measure residents' physical, social and cognitive function and clinical condition. Separate lists of covariates had been defined for CC and PAC QMs;

- indices and scales derived from the Resource Utilization Group-III (RUG-III) system, now used to adjust Medicare payments to nursing facilities -- these included the Nursing Case Mix Index (CMI), for both chronic and post acute care; scales created from the RUG CMI model (scales for Extensive Care, Clinically Complex, Cognitive Impairment, Extensive Services, Late Loss ADL, Behavior Problems, Rehabilitation and Special Care); and the Cognitive Performance Scale (CPS);

- the Nursing Severity Index (NSI), both weighted and unweighted;

- an "end of life" measure, the Personal Severity Index (PSI), and two subcomponents: PSIS1 that captures clinical indicators, and PSIS2 that includes functional indicators;

- a Cardio/Pulmonary Impairment Severity Scale;

- twelve diagnosis indicators, from items in Sections I and J of the MDS.

Exhibit 2 lists all the covariates that the SC proposed for testing at the September meeting. Starred items in Exhibit 2 were eventually dropped from the list of recommended covariates. Attachment 1 in the Technical Appendix presents specifications for the all the covariates that we tested.

Step Two: Reviewing Correlation of Covariates and QMs at the Resident Level

To be useful for adjustment, covariate measures have to correlate with the QMs. Using our previous example, unless the prevalence of a proposed incontinence covariate increases as the prevalence of pressure sores increases, incontinence will be a poor covariate choice for the pressure sore QM. Note that the direction of correlation need not be positive. More residents needing help in bed mobility should be associated with fewer residents showing improvement in walking. Measuring the strength and direction of the correlation between QMs and proposed covariates was the next step in our analysis.

Following the September meeting, SC investigators calculated statistical measures of correlation of all proposed covariates with the 45 QMs. In the following discussion, these correlation measures are referred to collectively as "R-statistics."[5] Attachments 2 and 3 display these statistics for all covariates and CC and PAC QMs respectively.

The SC used these statistics to refine the list of covariates for succeeding phases of the analyses. In general, we judged each correlation against a common threshold: if the R-statistic was greater than 0.10, we retained the covariate for further testing. However, in some instances in which clinical and other factors seemed compelling, the SC retained covariates with R-statistics lower than 0.10. In other cases where statistical criteria were met, we rejected covariates that we believed failed to meet other criteria for validity.

Original Covariates

Covariates for chronic care QMs. Most of the original covariates showed correlation with one or more QMs, but not always the QMs to which they had originally been linked. For 1,140 QM/covariate relationships, 161 R-statistics exceeded 0.10, and 89 exceeded 0.15. Of the 30 covariates that had been included in CC QM specifications, only six achieved R-statistics of 0.10 or greater. For several QMs, there were no covariates that made the cut, including QMs with and without resident-level covariates in the current specifications. However, some covariates seemed to be closely associated with several QMs. Only one covariate (age greater than 76) had no apparent association with any QM, though five others had only one (unsteady gait, falls in past 30 days, planned discharge, ALS or MS diagnosis, and unstable function). "Large" R-statistics tended to cluster in association with particular QMs or groups of QMs. This was most obvious for the incontinence QMs, with 13 R-statistics of 0.40 or higher. Directional signs for most of the measures that made the cut seemed appropriate. For example, wandering behavior was positively correlated with the QM Prevalence of Antipsychotic Drug Use (CDRG01).

Other examples included the following:

- The QM ADL Worsening Following Improvement (CADL01) had no covariates in the original specifications, and analyses confirmed no association with any of the candidate covariates.

[5] At the resident level, QMs are measured as "dichotomous variables" (a resident either has or does not have a problem), and most of the proposed covariates are similarly dichotomous. The SC used the phi statistic as a measure of correlation between two dichotomous variables. As with most measures of correlation, the phi statistic ranges from –1 to +1. The larger the absolute value, the closer the estimated association between a covariate and a problem. Phi's of +. 32 and -.32 show equally strong association, but the positive statistic implies that the covariate and QM move together in the same direction (larger values of the covariate are associated with larger values of the QM), while a negative statistic shows the two moving in opposite directions. For covariates (like the CMI) that could take on many values, we used the coefficient of determination, computed from a logistic regression, as a measure of QM/covariate correlation.

- For the Cognition Worsening QM (CCOG01), none of the four original covariates (bowel incontinence, fall in the past 30 days, weight loss, and age greater than 76) reached the 0.10 threshold.

- The Worsening Bowel Incontinence QM (CCNT02) had three original covariates (short-term memory problem, bladder incontinence, and dependence in dressing), each of which was confirmed in statistical tests, along with another covariate not in the original list (long-term memory problem).

Covariates for PAC QMs. Of the 56 PAC/covariate R-statistics, eleven exceeded 0.10, and seven exceeded 0.15. Four of these did so in association with the QMs they were originally designed to adjust.

Specific examples included:

- The PAC Pressure Sore QM (PPRUX01)had five covariates (sore resolved, needs bed mobility assistance, bowel incontinence, diabetes/peripheral vascular disease and low body mass index); statistical analysis showed correlation with this QM for the first three.

- The Improvement in Walking QM (PWAL0X) had no original covariates, but statistical analysis confirmed correlation with three (no prior residential history, needs bed mobility assistance and bowel incontinence).

RUG-III Case Mix Index (CMI)

The RUG-III Nursing Case Mix Index (CMI) was computed from the 44-category RUG-III Grouper, separately for CC and PAC residents. The CC CMI as a covariate performed well for some QMs. Thirteen of a total 30 QMs that were tested for correlation with the CMI had R-statistics that exceeded 0.10 (nine of the 13 exceeded 0.15). High R-statistics were concentrated in pressure sore and incontinence QMs. For the PAC QMs, the CMI correlated with four of seven QMs at an acceptable level (three of the four R-statistics exceed 0.15). Directional signs all seemed appropriate for both CC and PAC QMs: positive for QMs where higher numbers imply worse performance, and negative for the few cases, such as Improvement in Walking (CWAL0X and PWAL0X), where the positive signs mean improvement.

RUG Scales

We tested seven scales generated by the RUG-III Grouper as intermediate steps in constructing the CMI, for both CC and PAC residents. [6] These scales included:

- Late Loss ADL,
- Behavior Problems,
- Clinically Complex,
- Extensive Care,

[6] To test these scales, we converted several to dichotomous (two-valued) measures. Specifications for recoding may be found in Attachments 6 and 7.

- Cognitive Impairment,
- Rehabilitation, and
- Special Care.

Late Loss ADL achieved R-statistics at an acceptable level for the most CC QMs (11 out of 38) and PAC QMs (three out of seven). Other scales ranged in the number of acceptable R-statistics from five to eight for CC QMs and from none to two for PAC QMs.

Nursing Severity Index (NSI)

The Nursing Severity Index (NSI) is based on the presence or absence of 30 nursing diagnoses that researchers developed to predict morbidity, mortality and length of stay in acute care residents. For testing as a potential covariate, we measured the NSI diagnoses from items in the MDS, with most relying on the presence or absence of multiple MDS items. For example, the nursing diagnosis "less nutrition than required" is in effect if the MDS shows weight loss, parenteral/IV, feeding tube, or nutrition/hydration to manage skin problems. For each resident, the unweighted NSI is a number between 0 and 30 that represents the sum of all nursing diagnoses coded for that resident. We also tested a version of the NSI in which the components were weighted by the inverse of each component's frequency in the total resident population. The assumption was that rare components represent relatively high levels of acuity and resource intensity, and should thus have more weight than more frequent, but less serious, components. We analyzed the unweighted and weighted NSIs, for both CC and PAC residents.

Weighting the NSI made a difference, but not an improvement, over the unweighted NSI. The weighted NSI that was applied to CC QMs performed poorly overall. R-statistics equaled or exceeded 0.10 for only three of 38 QMs. R-statistics were acceptable for the incontinence (PCNT0X) and pressure sore (PPRU0X) PAC QMs. Unweighted, the NSI produced 23 R-statistics meeting or exceeding the 0.10 threshold (17 equaled or exceeded 0.15). The unweighted NSI was closely associated with six of the seven PAC QMs, all with R-statistics that exceeded 0.15.

Personal Severity Index (PSI)

The Personal Severity Index (PSI), developed by investigators at Hebrew Rehabilitation Center for Aged and the University of Michigan, is designed to capture functional and clinical conditions characteristic of residents who are near death.[7] (See Attachment 13 in the Technical Appendix) We tested both the full PSI, with 18 components derived from MDS data and age (90 or older), and two subcomponents: PSIS1 (eight clinical components plus age) and PSIS2 (nine functional components plus age). In general, all three versions of the PSI performed about equally well (10 or 11 R-statistics 0.10 or higher for CC QMs, and two to four for PAC QMs).

[7] PSI development and validation is described in Morris, J.N., R. Jones, S. Morris and B. Fries. Proximity to Death, a Modeling Tool for Use in Nursing Homes. Hebrew Rehabilitation Center for Aged. December 2002. (Attachment 13)

Other Scales

Cognitive Performance Scale. The Cognitive Performance Scale (CPS) is a six-value measure derived from the RUG Grouper. For CC QMs, the CPS achieved about the same frequency of acceptable correlations as the PSI (10 out of 38). The scale was correlated with five of seven PAC QMs.

Cardio/Pulmonary Impairment Severity Scale. This scale failed to demonstrate any association with any of the CC or PAC QMs, coming closest (with an R-statistic of –0.09) for the PAC Respiratory Problems QM (PRSP0X).

MDS Diagnosis Indicators

In addition to one original covariate (diagnosis of ALS or MS), we tested 12 new diagnosis indicators from Sections I and J of the MDS. These included:

- acute episode or flare-up,
- Alzheimer's disease,
- dementia other than Alzheimer's,
- arteriosclerotic heart disease (AHSD),
- arthritis,
- cancer,
- congestive heart failure (CHF),
- depression,
- emphysema/COPD,
- hip fracture,
- osteoporosis, and
- renal failure.

Of the 12, only acute episode or flare-up (with two R-statistics of 0.10 or higher), Alzheimer's disease (three) and dementia other than Alzheimer's (four) showed any association with any of the chronic care QMs. Alzheimer's and other dementia achieved accepted R-statistics in relation to three PAC QMs, as did hip fracture and arthritis, both in relation to the inadequate pain management PAC QM. In the final specifications, we combined Alzheimer's disease and dementia other than Alzheimer's into one covariate.

Selecting Covariates for Further Testing

At this and every stage of the analysis/decision process, the SC reviewed the clinical and behavioral justification for certain covariates, along with the growing body of evidence on statistical relationships. Before Step Three, we eliminated 17 of the original 30 covariates, RUG-based scales for Extensive Care, Extensive Services, Rehabilitation and Special Care, the NSI, the Cardio/Pulmonary Impairment Severity Scale, and all diagnosis indicators except acute episode or flare-up and Alzheimer's disease and other dementia. In general, decisions to drop the original covariates were based on statistical evidence and concerns about over adjustment. We dropped the Cardio/Pulmonary Impairment Severity Scale because it failed every statistical test. Both versions of the NSI were dropped because, though the unweighted version performed well statistically, the SC came to see the NSI as including too many of the components that were seen to threaten over

adjustment in the original covariates. Discarded RUG scales were those that the SC believed to be excessively gameable, because their definitions depended heavily on levels of service. Finally, all diagnostic indicators that were dropped fell short of the 0.10 correlation threshold.

Step Three: Testing Covariates in Resident-Level Prediction Models

Although the SC recommended no covariates or only one covariate for some QMs, for others there were several candidates. We took the position that a valid method for adjusting QMs could include several separate measures of risk. However, we needed to understand how each proposed covariate performed in concert with other covariates. Covariates may be correlated with each other as well as with QMs. For example, the QM Worsening Bowel Incontinence (CCNT02) had one-to-one correlation with two covariates: the Late Loss ADL scale (R = 0.15) and dependence in dressing (R = 0.12). But Late Loss ADL and dependence in dressing themselves showed a high degree of correlation with each other (R = 0.69). So it was reasonable to suspect that these two covariates may not be independent measures of the risk of worsening bowel incontinence. In some cases, it may be appropriate to drop a covariate that may be duplicating the role of another covariate. Attachments 4 and 5 show measures of correlation between the proposed covariates.

To further the decision process on certain QMs with multiple proposed covariates, the SC examined the effects of all recommended covariates together. This process happened in two stages. First, we used a "test sample" of residents to build statistical models for each QM and its associated covariates. Second, we explored the capacity of these models to function as expected with entirely different "retest samples" of residents.

Constructing Prediction Models. The SC conducted initial analyses on 20 percent of the test sample residents. We used a multivariate statistical technique that related each resident's score on the problem captured by a QM (scored 1 if the problem was present, 0 if not present) to all the recommended covariates for that QM.[8] We assessed the contribution of each covariate to the whole group's correlation with a QM by adding covariates to the model one by one, usually beginning with the one that had shown the highest R-statistic, until all were included. Then we reversed the process, beginning with the covariate that had the lowest R-statistic. We used statistics from this process to determine whether or not covariates continued to demonstrate an association with the QM even after other covariates were entered, and to determine which combinations of covariates most accurately predicted the frequency of resident problems. Measures of how well the model "fit" actual resident data included a multiple-measure version of the R-statistic, and measures that captured how frequently the combined covariates predicted a resident's status correctly. These included "concordance," an overall measure of the percent of residents with and without the problem that the model predicted correctly, and "sensitivity," the percent of residents with the problem that were predicted to have the problem.

Taking these measures into account, the SC made recommendations for covariates to be included in the final resident-level models, constructed on data from the entire test sample.

[8] For these analyses, the SC used logistic regression to conduct all multivariate analyses.

Testing the prediction models. Next, we used the "retest sample" to see if the test models built in Step Three would perform equally well for an entirely different set of residents. To do this, we drew five 20 percent random subsamples of residents from the retest sample. Then, for each QM and each 20 percent sample, we used the test models to predict resident problems. We then studied the correlation of predicted problems with actual problems, using R-statistics, concordance and sensitivity measures. In general, the models performed equally well for the test sample and the retest subsamples. Measures of correlation and fit were not substantially different – that is, the retests did not improve on results from the test models, but they also did not show poorer performance. Test models, and the retest results, are reported in Attachment 8.

Preparing for the facility level analyses. The SC reviewed results of the retest, and revisited issues of clinical validity for several of the proposed covariates. At the conclusion of this process, the SC confirmed the selection of covariates for further testing at the facility level.

Step Four: Analyzing Facility-Level Risk Adjustment

Although the basis for a valid risk adjustment system is correlation between covariates and QMs at the resident level, QMs are risk adjusted at the facility level. In other words, risk adjustment is performed for all residents with a specific condition or set of conditions across the whole facility, rather than on a resident-by-resident basis. In the next stage of the analyses, we used our models to risk adjust facility QMs. We then compared distributions of QMs with and without adjustment, to assess what impact adjustment had on facility scores and rankings.

Computing Adjusted QMs

For QMs with recommended covariates, we computed adjusted scores for all 16,615 U.S. nursing facilities in Q2 2002, applying the quality measure calculation method currently in use for the 10 publicly reported QMs. There are three components of an *adjusted facility QM score*.

1. The *observed facility QM score* is simply the ratio of a facility's residents who have the condition or problem represented by the QM over all facility residents at risk for the problem.

2. The *predicted facility QM score* is the score that would be predicted for the facility, given the mix of residents residing in the facility for the time period under analysis. We predicted each facility's score by combining the appropriate covariates and profiles of resident risks characteristics, using models developed in Step Three (above).

3. The *national average QM score* completes the formula.

Each facility's adjusted score combines the observed score, the predicted score and the national average score. For the publicly reported QMs, no adjusted scores are computed for facilities with at-risk resident populations smaller than a set threshold (30 residents for CC QMs, and 20 for PAC QMs). This is because QM scores based on small numbers of residents tend to be highly unstable, making them difficult to interpret. (An increase from one to two residents with a QM problem, out of a pool of five residents at risk for the problem, will double the QM score from 20 to 40 percent). The SC adhered to this exclusion rule for adjustment based on the new covariate models.

Assessing the Effects of Adjustment: Comparing the New Models to Earlier Models

We assessed the effects of adjustment on facilities in several ways. Wherever possible, we compared performance of the new models to alternative models. These alternatives were based on the facility admissions profile (FAP), on the original resident level covariates, or on FAP/original covariate combinations. First, we assessed the overall impact of adjustment, measured through correlation (of observed and adjusted QMs) and through analysis of how facility ranking on QM scores changed after adjustment. Then we studied more closely the extent to which different adjustment methods targeted facilities that admit residents with more or less risk.

Correlating observed and adjusted QMs. If an observed QM correlates less closely with QMs adjusted with the new covariates than QMs adjusted by other methods, then the new covariate model "works" in the sense that the new model changes facilities' QM scores more than the other models. We correlated the observed QMs with QMs adjusted by the new covariates. Where appropriate, we did the same for FAP-adjusted and original covariate-adjusted QMs. Results are presented in Attachment 9. We made several observations based on these analyses.

- Measures of correlation between observed QMs and QMs adjusted by the new covariates ranged from 0.986 (very close correlation, little impact of adjustment on the QMs) to 0.713 (less close correlation, more impact).

- For half of the 40 QMs that had alternative adjustment models, the new covariate models had more impact (lower correlation with observed QMs) than at least one of the alternatives. For 14 QMs, one of the alternatives was the original covariate model. The new covariate models showed greater impact (lower R) than the original covariate model for 10 of these QMs. Only for Worsening Behavioral Symptoms (CBEH04), Worsening Bowel Incontinence (CCNT02), Worsening Bladder Incontinence (CCNT03) and Worsening Pressure Sores (CPRU04) did the original covariate models appear to have greater impact than the new models.

- In general, models based on the FAP showed lower correlation than the new covariate models. For 18 QMs, the FAP model was the only alternative. Of these 18, the new models showed lower correlation than the FAP for only six.

- We proposed new covariate models for two QMs that had no previous adjustments: ADL Improvement (CADL03) and Little or No Activity (CSOC02). However, measures of correlation with observed QMs were high for both (0.961 and 0.984 respectively), suggesting minimal impact.

Exploring the effects of adjustment on facility ranking. Correlation provides a good measure of average impact, but it does not show the effects adjustment might have on facility rankings by QM score. We began our exploration of ranking by documenting overall changes, in terms of movements among facilities grouped into the highest and lowest 10 percent of QM scores.

To study the effects of adjustment on facility ranking, we computed several measures, including

- the percent of facilities ranked in the highest and lowest 10 percent of each QM distribution that moved toward the "middle" of the distribution after adjustment (from the highest 10

percent to any lower rank above the mid-point of the distribution, and from the lowest 10 percent to any higher rank below the mid-point) — for this measure, the larger the percent of facilities moved from the highest and lowest groups toward the middle groups, the more "effective" the adjustment method;

- the percent of facilities that did not move at all — here, the smaller the percent of non-movers, the more effective the adjustment method;

- the average difference in percentage points between the observed and the adjusted QM scores — the larger the average difference between observed and adjusted QMs, the more effective the adjustment method.

In general, as Attachment 9 shows, statistics on changes in rank confirmed evidence from the correlation analyses. For the 10 QMs for which correlation analysis showed the new models to be more effective, there was more movement out of the top and bottom 10 percent of QM scores and fewer non-mover facilities for the new than for the original covariate models. For the QM Prevalence of Indwelling Catheter (CCAT02), scores for 41 percent of facilities were unchanged by adjustment under the new model (Line 3, Column 3), compared to 47 percent under the original covariate model (Line 3, Column 2). Adjusted with the new model, 3.79 percent of facilities in the top 10 percent and 1.26 percent in the bottom 10 percent (Lines 1 and 2, Column 3) moved toward the middle (compared to 2.86 percent and 0.99 percent respectively for the original covariate model, shown in Lines 1 and 2, Column 2). The average percentage point change due to adjustment was 0.0191 for the new model and 0.0162 for the original (Line 1, Columns 2 and 3).

In contrast, for the QM Worsening Behavioral Symptoms (CBEH04), adjustment using the new covariates left 68.8 percent of facilities unchanged in ranking, compared to 64.9 percent of facilities under the original covariate model. With the new model, movement from the top 10 percent (1.25 percent) and bottom 10 percent (0.60 percent) was less than movement with the original covariate model (1.50 and 0.82 percent respectively).

Taken together, the correlation and rank analyses do not show many clear patterns. However, we found some interesting tendencies.

- There are three prevalence QMs for pressure sores (CPRU01 - CPRU03), and three for incontinence (CCNT01, CCNT05 and CCNT06). For each, the new models outperformed the FAP models and, when they were present, the original covariate models. Exactly the opposite was true for the incidence measures Worsening Pressure Sores (CPRU04), Worsening Bowel Incontinence (CCNT02), and Worsening Bladder Incontinence (CCNT03), for which the new models proved to be less effective than their alternatives.

- For all PAC QMs, the FAP models generated more movement than the new models. But for three PAC QMs, Failure to Improve During Early PAC Period (PADL0X), Failure to Prevent or Improve Pressure Sores (PPRU0X), and Failure to Prevent or Improve Respiratory Problems (PRSP0X), the new covariate models were more effective than the original covariate models.

Exploring the effects of adjustment on facilities admitting high and low-risk residents. To this point, we can only say that one adjustment model had a greater impact than another either on the whole distribution of QM scores or on facilities with extremely high or low scores. We do not know how effectively any adjustment performed in targeting QMs of facilities that care for very high or low risk residents.

To study targeting effectiveness, we created facility-level indicators of risk from the RUG-III CMI scores of residents at admission. We classified facilities as "High QM/High CMI" if they were in the top 10 percent of the QM distribution and the top 10 percent of the CMI distribution. "Low QM/Low CMI" defined groups of facilities in the bottom 10 percent of both distributions. Then we computed two targeting ratios: High QM/High CMI facilities, as a percent of all facilities that moved from the top 10 percent of a QM distribution, and equivalent ratios for Low QM /Low CMI facilities moving from the bottom 10 percent of a QM distribution. If the new models targeted for risk more effectively than the alternatives, we should expect to see larger targeting ratios for the new models.

Overall, as Attachment 10 shows, the new covariate models did a better job of targeting than the original covariate models in 18 of 28 possible comparisons (comparing movement from the top and bottom 10 percent in each of 14 QMs). An example will show how we reached these conclusions. For Worsening Behavioral Symptoms (CBEH04), we first calculated the percent of all 16,615 facilities that moved from the top and bottom 10 percent for this QM. Column 5 shows that 0.816 percent (136 facilities) moved from the bottom 10 percent to the middle of the distribution when adjusted using the original covariate model, while 0.602 percent (100 facilities) moved when adjusted by the new model (Column 6). Some of these "movers" were in the lowest CMI risk group. Using the original covariate model, 0.123 percent of all facilities (20 facilities) that moved were low in both QM and risk (Column 2). For the new model, the percentage was 0.122 (19 facilities), shown in Column 4. The percent of "movers" from the low QM group that were also low risk, shown in Column 7, was about 15 percent (0.123/0.816) for the original covariate model. The corresponding figure for the new model was about 20.2 percent (0.122/0.602), shown in Column 9.

For seven QMs, the new models were superior at both extremes of the distribution. For four more, the new model outperformed the original model at one or the other extreme, but not both. For three, the original covariate model targeted facilities better than the new model.

The new covariate models also performed quite well in comparison to FAP-adjusted models, shown in Columns 8 and 11. Out of 56 possible comparisons (comparing movement from the top and bottom groups in each of 28 QMs), the new models were more effective than FAP-adjusted models in 35 comparisons.

Adjustment and QM Change at the Facility Level -- Seven "Case Studies"

To provide concrete illustrations of how adjustment affects QM scores, we tabulated 40 facility-level measures for each of seven QMs. For this purpose, we selected samples of 20 facilities from the High QM/High CMI group, and 20 from the Low QM/Low CMI group for the following QMs (starred QMs are publicly reported):

Prevalence of Infections (CINF0X)*
Pressure Sore Prevalence (high and low risk) (CPRU01)*
Inadequate Pain Management (CPAI0X)*

Inadequate Pain Management (PPAI0X)*
Improvement in Walking (PWAL0X)*
Prevalence of Feeding Tubes (CNUT01)
Worsening Bladder Incontinence (CCNT03)

Attachment 11 shows how the new covariate models adjusted each facility's observed QM, compared to the FAP-adjusted and original covariate models. In most cases, adjustment reduced the scores of the High CMI facilities and increased scores of the Low CMI facilities. For example, adjusted by the new covariate model, High CMI Facility 12's QM for CINF0X dropped from 30.26 percent to 25.79 percent. For Low CMI Facility 17, their CINF0X QM increased when adjusted by the new model from 4.00 percent to 4.65 percent. We did not find entirely consistent patterns of increase or decrease across all facilities. This was the expected result. Since the CMI is an inclusive measure of resident case mix, we would not expect to see adjustment models achieve uniformly high targeting efficiency across diverse facilities.

Assessing the Effects of Adjustment: Comparing the New Models Across QMs

We compared the new covariate models across QMs, to give an additional perspective on relative performance. This also provided context for assessing QMs that had few (or no) alternative adjustment models for comparison. Attachment 12 ranks all QMs in ascending order on four measures of movement following adjustment. In this table, Column 1 shows the percent of all facilities that moved from the lowest 10 percent of observed QM scores. Column 2 shows the percent moving from the highest 10 percent. Columns 3 and 4 report "targeting ratios." Column 3 shows the lowest-CMI facilities that moved, as a percent of all movers from the lowest 10 percent of QM scores. Column 4 shows the highest-CMI movers, as a percent of all movers from the highest 10 percent of QM scores. QMs in each column are ranked above and below the median value for each measure. We considered new covariate models above the median to be "relatively effective," compared to models for QMs below the median.

Here, as in the earlier comparisons among alternative adjustors, we were particularly interested in targeting capability, shown in Columns 3 and 4. The range for models targeting low risk facilities runs from zero for Antipsychotic Use, High and Low Risk (CDRG01) -- meaning that this model moved none of the lowest risk facilities toward the middle of the distribution -- to 53.4 percent (over half of facilities that moved from the lowest scores for Worsening Pressure Sores, CRPU04, were from the lowest risk group). At the other extreme, the Walking Improvement (PWAL0X) model did not move any high-risk facilities out of the highest group of QM scores, while about 35 percent of CPRU04's high-QM movers were high-risk facilities.

Some QMs performed consistently above the median on all four measures, while others changed their relative positions. The former group included CINF0X, CCNT01, CCNT02, CCNT06, CPRU01, CNUT01, PPRU0X and PCNT0X. Others were above the median in both measures of targeting effectiveness, but below on one or more measures that captured total movement: these include CADL03, CCAT02, CMOB1, CCNT03 and CPRU04. From these results, it is apparent that incontinence and pressure sore QM models tended to be relatively effective at targeting facilities based on risk.

Summary of New Covariates Performance at the Facility Level

We can summarize with a few generalizations about performance of the new adjustment models.

- Taken as a group, the new models tended to do a better job of targeting adjustment appropriately. That is, the new models moved facilities ranked highest both on QM scores and resident risk toward the center of the distribution, and did the same for facilities ranked lowest on QM and risk. This was particularly true in comparisons of new to original covariate models, but it was also true in many comparisons with FAP adjustment models.

- Generalizations about the performance of new models for QM "families" are difficult to support. But the new covariate models for CC and PAC incontinence and pressure sore QMs tended to perform well, compared to FAP and to original covariate alternatives, and compared to new covariate models for other QMs.

- As expected, the performance of some QM models did not conform neatly to any of these generalizations. These included the two QMs for which we have only new covariate models, ADL Improvement (CADL03) and Little or No Activity (CSOC02). For these, we can only cite their performance relative to other new covariate models. CADL03's model performed above the median in targeting effectiveness for both high and low risk facilities. CSOC02 only performed above the median in the percent of facilities moved from the lowest QM group.

  For six QMs (CCOM01, CDRG01, CDRG02, CPRU02, CFAL01, PPAI0X and PWAL0X), the new models were less effective in targeting than the alternatives in every comparison, though some showed effectiveness in other ways. Only PWAL0X failed to perform well in comparisons with the alternative (FAP-adjustment) in total and targeted movement effectiveness, and in comparison with other new covariate-adjusted QMs.

Conclusions and Recommendations

This report describes a process through which the SC analyzed potential covariates for adjusting nursing facility CC and PAC QMs. We recommend resident-level covariates and models, shown in Exhibit 3 that represent a significant departure from the original QM specifications.

- None of the recommended models includes a facility admissions profile (FAP) measure.

- Some QMs with models based on original covariates have entirely new covariates, or no recommended covariates at all.

- Some QMs formerly adjusted only with FAP measures now have models based on resident-level covariates.

- Some QMs, for which no adjustment model was recommended, now have new covariate models.

The SC began with an inclusive list of potential resident-level covariates. We tested the correlation of these covariates, singly and jointly, with the 45 QMs. We reviewed each covariate against standards of clinical validity in the context of particular QM models. We tried to follow a "conservative" strategy designed to minimize the chances of over adjustment.

In general, facility-level analyses of adjustment based on the new covariate models showed improvement over the FAP and/or original covariate models. We did not, however, use results from the facility-level analyses to revise our recommendations on new covariate models. As expected, we found variation in comparative performance. Some new models achieved solid improvement measured in almost all comparisons with alternatives, most performed well in some comparisons, and a few showed little or no relative improvement. These results suggest that further exploration of the role of facility-level measures may produce more effective adjustment models. But we consider the resident-level covariates that we recommend to be valid and accurate for adjusting publicly reported nursing facility QMs.

Exhibit 1 – Table of QMs
*** QMs selected for public reporting**

Chronic Care Quality Measures	Code Name
* Percent of residents who had an unexpected loss of function in some basic daily activities	CADL01
Percent of residents with worsening function in some basic daily activities	CADL02
Percent of residents who have improved in their ability to function	CADL03
Percent of residents who have declined in their ability to locomote	CMOB01
Percent of residents who walk as well or better than the previous assessment	CWAL0X
Percent of residents whose cognitive ability has worsened	CCOG01
Percent of residents whose ability to communicate has worsened	CCOM01
Percent of residents with symptoms of delirium	CDEL0X
Percent of residents with inappropriate behavior (high & low risk)	CBEH01
Percent of residents with inappropriate behavior (high risk)	CBEH02
Percent of residents with inappropriate behavior (low risk)	CBEH03
Percent of residents whose behavior has worsened	CBEH04
Percent of residents who have become more depressed or anxious	CMOD03
Percent of residents engaging in little or no activity	CSOC02
Percent of residents with a new indwelling catheter	CCAT01
Percent of residents with indwelling catheters	CCAT02
Percent of residents who are bladder or bowel incontinent (high & low risk)	CCNT01
Percent of residents who are bladder or bowel incontinent (high risk)	CCNT05
Percent of residents who are bladder or bowel incontinent (low risk)	CCNT06
Percent of residents with worsening bowel continence	CCNT02
Percent of residents with worsening bladder continence	CCNT03
Percent of residents with a urinary tract infection	CCNT04
Percent of residents who have fallen	CFAL01
* Percent of residents with infections	CINF0X
Percent of residents with a feeding tube	CNUT01
Percent of residents with a low BMI	CBMI0X
Percent of residents who have unexplained weight loss	CWGT01
* Percent of residents with pain	CPAI0X
Percent of residents with worsening pain	CPAN01
* Percent of residents with pressure sores (high&low risk)	CPRU01
Percent of residents with pressure sores (high risk)	CPRU02
Percent of residents with pressure sores (low risk)	CPRU03
Percent of residents with worsening pressure sores	CPRU04
Percent of residents with burns, skin tears or cuts	CBUR0X
* Percent of residents in physical restraints	CRES01
Percent of residents on antipsychotics without a diagnosis of psychosis (high & low risk)	CDRG01
Percent of residents on antipsychotics without a diagnosis of psychosis (high risk)	CDRG02
Percent of residents on antipsychotics without a diagnosis of psychosis (low risk)	CDRG03

Post Acute Quality Measures	Code Name
* Percent of short-stay residents with delirium	PDEL0X
Percent of short-stay residents who have not improved since admission	PADL0X
Percent of short-stay residents whose ability to control their bowel or bladder has not improved since admission	PCNT0X
* Percent of short-stay residents with pain	PPAI0X
Percent of short-stay residents whose pressure sores have not gotten better	PPRU0X
Percent of short-stay residents who have developed a respiratory infection or have not gotten better	PRSP0X
* Percent of short-stay residents who walk as well or better on day 14 as on day 5 of their stay	PWAL0X

Exhibit 2 – Table of Covariates[9]

CHRONIC CARE	
COVARIATE NAME	**CODE NAME**
Age > 76*	CCOG1_D
ALS/MS diagnosis	CNUT1_B
Any wandering	CFAL1_B
Bed mobility problem*	CWGT1_B
Bladder Incontinence*	CCNT2_C
Bowel incontinence*	CCOG1_A
Fall in last 30 days*	CCOG1_B
Fall in last 180 days*	CMOB1_A
Independence in daily decision making*	CPAIX_A
Locomotion Problem	CPRU4_D
Long term memory problem	CWGT1_A
Moderate/impaired decision making problem	CBEH4_B
Modes of expression: speech*	CBEH4_A
More dependence in dressing*	CCNT3_B
More dependence in toileting	CMOB1_C
Motor agitation	CBEH4_C
Not totally dependent in transferring	CMOD3_A
Pain Present*	CMOD3_B
Physically abusive behavior*	CWGT1_C
Planned discharge: 30-90 days*	CMOD3_C
Pressure sores (stage 3 or 4)	CCAT1_B
Requires much assistance for eating	CCOM1_A
Resident not bedfast*	CFAL1_A
Severe decision making problem*	CCNT3_C
Short term memory problem	CCOM1_B
Swallowing problem	CNUT1_A
Transferring problem*	CPRU4_A
Unstable condition*	CPRU4_B
Unsteady gait/cognitive impairment	CFAL1_C
Weight loss (5%, past 30 days; 10%, past 180 days)*	CCNT3_D
Nursing Severity Index (NSI)	
Weighted NSI*	NSIDX
Unweighted NSI*	NSIUNWT
Personal Severity Index (PSI)	
Full PSI	MFI
PSI: Subset 1 – Diagnoses	MFIS1
PSI: Subset 2 – Non-Diagnoses	MFIS2

[9] Starred items were dropped from final QM specifications, either because they failed the SC's statistical threshold or because literature and informed judgment of SC members determined that they failed on non-statistical grounds (over adjustment, clinical relevance, potential for gaming, etc.)

CHRONIC CARE	
COVARIATE NAME	**CODE NAME**
Resource Utilization Group (RUG)	
RUG Nursing CMI	R_CMIC
RUG Late Loss ADL	R_ADL
RUG Behavior Problems	R_BEH
RUG Clinically Complex	R_CLN
RUG Extensive Care*	R_EXT
RUG Cognitive Impairment	R_IMP
RUG Rehabilitation*	R_REHC
RUG Special Care*	R_SPC
Other Scales	
Cardiopulmonary Severity Scale*	CARDIO
Cognitive Performance Scale (CPS)	CPS
MDS Diagnosis Indicators	
Acute Episode or Flare-up	J5B
Combination Alzheimer's Disease / Other Dementia	I1QU
Arteriosclerotic Heart Disease (AHSD)*	I1D
Arthritis*	I1L
Cancer*	I1PP
Congestive Heart Failure (CHF)*	I1F
Depression*	I1EE
Emphysema/COPD*	I1II
Hip Fracture in last 180 days*	J4C
Osteoporosis*	I1O
Renal Failure*	I1QQ

POST-ACUTE CARE	
COVARIATE NAME	**CODE NAME**
Bowel incontinence*	PPRUX_C
Diabetes or peripheral vascular disease*	PPRUX_D
Indicator of asthma on prior assessment*	PRSPX_A
Indicator of emphysema/COPD on prior assessment*	PRSPX_B
Low body mass index*	PPRUX_E
Needs bed mobility assistance*	PPRUX_B
No prior residential history*	PADLX_A
Sore resolved*	PPRUX_A
Nursing Severity Index (NSI)	
Weighted NSI*	NSIDX
Unweighted NSI*	NSIUNWT
Personal Severity Index (PSI)	
Full PSI	MFI
PSI: Subset 1 – Diagnoses	MFIS1
PSI: Subset 2 – Non-Diagnoses	MFIS2

POST-ACUTE CARE	
COVARIATE NAME	**CODE NAME**
Resource Utilization Group (RUG)	
RUG Nursing CMI	R_CMIP
RUG Late Loss ADL	R_ADL
RUG Behavior Problems*	R_BEH
RUG Clinically Complex	R_CLN
RUG Extensive Care*	R_EXT
RUG Cognitive Impairment*	R_IMP
RUG Rehabilitation*	R_REHC
RUG Special Care*	R_SPC
Other Scales	
Cardiopulmonary Severity Scale*	CARDIO
Cognitive Performance Scale (CPS)	CPS
Bowel incontinence*	PPRUX_C
Diabetes or peripheral vascular disease*	PPRUX_D
Indicator of asthma on prior assessment*	PRSPX_A
Indicator of emphysema/COPD on prior assessment*	PRSPX_B
Low body mass index*	PPRUX_E
Needs bed mobility assistance*	PPRUX_B
No prior residential history*	PADLX_A
Sore resolved*	PPRUX_A
Nursing Severity Index (NSI)	
Weighted NSI*	NSIDX
Unweighted NSI*	NSIUNWT
Personal Severity Index (PSI)	
Full PSI	MFI
PSI: Subset 1 – Diagnoses	MFIS1
PSI: Subset 2 – Non-Diagnoses	MFIS2
Resource Utilization Group (RUG)	
RUG Nursing CMI	R_CMIP
RUG Late Loss ADL	R_ADL
RUG Behavior Problems*	R_BEH
RUG Clinically Complex	R_CLN
RUG Extensive Care*	R_EXT
RUG Cognitive Impairment*	R_IMP
RUG Rehabilitation*	R_REHC
RUG Special Care*	R_SPC
Other Scales	
Cardiopulmonary Severity Scale*	CARDIO
Cognitive Performance Scale (CPS)	CPS
MDS Diagnosis Indicators	
Acute Episode or Flare-up*	J5B
Combination Alzheimer's Disease / Other Dementia	I1QU
Arteriosclerotic Heart Disease (AHSD)*	I1D
Arthritis*	I1L
Cancer*	I1PP

POST-ACUTE CARE	
COVARIATE NAME	**CODE NAME**
Congestive Heart Failure (CHF)*	I1F
Depression*	I1EE
Emphysema/COPD*	I1II
Hip Fracture in last 180 days	J4C
Osteoporosis*	I1O
Renal Failure*	I1QQ

Exhibit 3 – Table of Finalized QMs and Covariates
* QMs selected for public reporting

Chronic Care Quality Measures	Accepted Covariates
* Percent of residents who had an unexpected loss of function in some basic daily activities	NO COVARIATES
Percent of residents with worsening function in some basic daily activities	NO COVARIATES
Percent of residents who have improved in their ability to function	RUG Nursing CMI RUG Clinically Complex
Percent of residents who have declined in their ability to locomote	PSI: Subset 1 – Diagnoses More dependence in toileting Requires much assistance for eating
Percent of residents who walk as well or better than the previous assessment	Full PSI
Percent of residents whose cognitive ability has worsened	NO COVARIATES
Percent of residents whose ability to communicate has worsened	Short term memory problem Long term memory problem Cognitive Performance Scale
Percent of residents with symptoms of delirium	NO COVARIATES
Percent of residents with inappropriate behavior (high & low risk)	RUG Cognitive Impairment Cognitive Performance Scale Long term memory problem
Percent of residents with inappropriate behavior (high risk)	Long term memory problem Moderate/impaired decision making problem Any wandering
Percent of residents with inappropriate behavior (low risk)	Moderate/impaired decision making problem Cognitive Performance Scale
Percent of residents whose behavior has worsened	Moderate/impaired decision making problem Cognitive Performance Scale
Percent of residents who have become more depressed or anxious	NO COVARIATES
Percent of residents engaging in little or no activity	RUG Nursing CMI RUG Late Loss ADL
Percent of residents with a new indwelling catheter	NO COVARIATES
Percent of residents with indwelling catheters	Pressure sores (stage 3 or 4) RUG Nursing CMI RUG Clinically Complex ALS/MS diagnosis
Percent of residents who are bladder or bowel incontinent (high & low risk)	PSI: Subset 1 – Diagnoses PSI: Subset 2 – Non-Diagnoses RUG Late Loss ADL
Percent of residents who are bladder or bowel incontinent (high risk)	PSI: Subset 1 – Diagnoses PSI: Subset 2 – Non-Diagnoses RUG Late Loss ADL
Percent of residents who are bladder or bowel incontinent (low risk)	PSI: Subset 1 – Diagnoses PSI: Subset 2 – Non-Diagnoses RUG Late Loss ADL

Chronic Care Quality Measures	Accepted Covariates
Percent of residents with worsening bowel continence	RUG Nursing CMI PSI: Subset 1 – Diagnoses PSI: Subset 2 – Non-Diagnoses RUG Late Loss ADL
Percent of residents with worsening bladder continence	PSI: Subset 1 – Diagnoses PSI: Subset 2 – Non-Diagnoses Cognitive Performance Scale RUG Nursing CMI
Percent of residents with a urinary tract infection	NO COVARIATES
Percent of residents who have fallen	Locomotion Problem Not totally dependent in transferring Unsteady gait/cognitive impairment Any wandering RUG Late Loss ADL
* Percent of residents with infections	RUG Nursing CMI RUG Clinically Complex
Percent of residents with a feeding tube	RUG Clinically Complex Swallowing problem RUG Nursing CMI
Percent of residents with a low BMI	PSI: Subset 1 – Diagnoses
Percent of residents who have unexplained weight loss	PSI: Subset 1 – Diagnoses
* Percent of residents with pain	Cognitive Performance Scale Long term memory problem
Percent of residents with worsening pain	NO COVARIATES
* Percent of residents with pressure sores (high&low risk)	RUG Nursing CMI RUG Clinically Complex PSI: Subset 1 – Diagnoses PSI: Subset 2 – Non-Diagnoses RUG Late Loss ADL
Percent of residents with pressure sores (high risk)	RUG Nursing CMI RUG Clinically Complex PSI: Subset 1 – Diagnoses
Percent of residents with pressure sores (low risk)	RUG Nursing CMI
Percent of residents with worsening pressure sores	RUG Nursing CMI RUG Late Loss ADL
Percent of residents with burns, skin tears or cuts	NO COVARIATES
* Percent of residents in physical restraints	NO COVARIATES
Percent of residents on antipsychotics without a diagnosis of psychosis (high & low risk)	Motor agitation Moderate/impaired decision making problem RUG Behavior Problems RUG Cognitive Impairment Long term memory problem Cognitive Performance Scale Combination Alzheimer's Disease / Other Dementia
Percent of residents on antipsychotics without a diagnosis of psychosis (high risk)	RUG Behavior Problems RUG Cognitive Impairment

Chronic Care Quality Measures	Accepted Covariates
	Combination Alzheimer's Disease / Other Dementia
Percent of residents on antipsychotics without a diagnosis of psychosis (low risk)	RUG Behavior Problems RUG Cognitive Impairment Combination Alzheimer's Disease / Other Dementia Moderate/impaired decision making problem Motor agitation

Post Acute Quality Measures	Code Name
* Percent of short-stay residents with delirium	NO COVARIATES
Percent of short-stay residents who have not improved since admission	Cognitive Performance Scale PSI: Subset 2 – Non-Diagnoses
Percent of short-stay residents whose ability to control their bowel or bladder has not improved since admission	RUG Nursing CMI PSI: Subset 1 – Diagnoses PSI: Subset 2 – Non-Diagnoses RUG Late Loss ADL
* Percent of short-stay residents with pain	Cognitive Performance Scale Hip Fracture in last 180 days Combination Alzheimer's Disease / Other Dementia
Percent of short-stay residents whose pressure sores have not gotten better	RUG Clinically Complex PSI: Subset 1 – Diagnoses PSI: Subset 2 – Non-Diagnoses RUG Nursing CMI RUG Late Loss ADL
Percent of short-stay residents who have developed a respiratory infection or have not gotten better	RUG Clinically Complex
* Percent of short-stay residents who walk as well or better on day 14 as on day 5 of their stay	RUG Late Loss ADL PSI: Subset 2 – Non-Diagnoses RUG Nursing CMI Cognitive Performance Scale